

ADMINISTRATIVE INTERNAL USE ONLY

80-0103
4 FEB 1980

80-0283/L


MEMORANDUM FOR: Deputy Director for Administration

FROM: Charles A. Briggs
Inspector General

SUBJECT: Professional Applicant Test Battery

1. Forwarded herewith for your review and comment is a draft copy of Chapter IX, "Professional Applicant Test Battery," of the Inspector General's Report on the Agency's Recruitment System. The report of the psychological consultants, with executive summary and appendices, on which the draft chapter is based, is included for your information.

2. The draft and the consultants' report are not being disseminated elsewhere within the Agency pending receipt of your comments. Your submission of comments by 29 February 1980 would be most appreciated.


Charles A. Briggs

STATINTL

Enclosures,
As stated

ADMINISTRATIVE INTERNAL USE ONLY

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

IX. PROFESSIONAL APPLICANT TEST BATTERY (PATB)

✓ The PATB, an eight hour series of written tests, has been used by the Agency for more than 20 years as a tool for the selection of new professional employees. This inspection examined two aspects of the PATB -- its role in the processing and selection of new employees, and evidence concerning its reliability and validity. The first aspect was addressed by the inspection team, the second by two nationally recognized experts in psychological testing. The complete report prepared by the consultants appears as Appendix H to this report. Section A of this chapter summarizes the findings and conclusions of the inspection team concerning Agency use of the PATB as a selection device. Sections B, C, D, and E summarize the major findings and conclusions of the consultants. Section F presents the inspection team's assessment and recommendations.

✓ A. Role of PATB in Personnel Selection

Our findings are based on a questionnaire/survey of 900 employees entering on duty between 1 October 1977 and August 1979, and of 500 supervisors.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

We find that there is no consistent policy among Agency components with respect to use of the PATB. The test battery is not administered to all applicants for professional positions, and, when it is administered, the results are not used uniformly. Among professional employees, two-thirds report they had taken PATB and one-third report that they had not.

More than 60% of the supervisors either have no opinion about PATB's usefulness or indicate that it is not used by their components. Only about one-fourth of the supervisors indicate that they give significant weight to the applicant's performance on PATB in making employment decisions. The more senior the supervisor, the longer the time of service or supervisory experience of the respondents, the less the value placed on PATB's usefulness.

Apart from the PATB, some Agency units are using tests constructed in the units to select professional personnel. There is no evidence that these tests have been validated, nor are validation data available for other selection tools, such as the job interview or transcripts.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

ADMINISTRATIVE-INTERNAL USE ONLY

Such inconsistencies are serious not only on grounds of fairness or legality, but are unwise in terms of Agency goals in hiring the most qualified personnel.

There is a need for the Agency to develop a systematic policy on personnel selection practices and the role of PATB in personnel selection.

In particular, policy guidelines need to be set as to which applicants for professional jobs should be required to take PATB, and such guidelines should be followed by all components. If the decision to require or not require PATB for professional jobs is made by individual managers there is a high potential for violating U.S. Equal Employment Opportunity Commission (EEOC) Guidelines on disparate treatment of applicants. In the opinion of the Office of General Counsel, "If CIA does not comply with lawful and appropriate EEOC regulatory issuances, CIA would stand in violation of both statute and executive order."^{1/}

1/

Memorandum from the Deputy General Counsel to the Director of Equal Employment Opportunity, OGC 79-05429, 13 June 1979.

ADMINISTRATIVE-INTERNAL USE ONLY

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

There is need also to validate all tests and procedures used for the selection of personnel and to control the use of unvalidated personnel selection procedures. We believe senior management should start an Agency-wide job analysis as the basis for a sound test development program.

B. Reliability

A fundamental consideration in evaluating a test battery is that of determining the reliability^{2/} of the tests which comprise the battery.

If test scores are unreliable, the scores will not predict anything. Fragmentary reliability data are available from research performed in the mid-1950's, but since 1958 only one reliability study has been done at the Agency. None of the available evidence is encouraging.

2/

A measurement procedure is reliable to the extent that repeated measurement gives consistent results for the individual - consistent in that his score remains substantially the same when measurement is repeated or in that his standing in the group shows little change.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

For white males and females, only five of the 31 PATB test scores have reliabilities that are at least minimally acceptable for making decisions about individuals. For all other tests and scales in PATB for which reliability data are available, the reliabilities are too low; thus the scores from them are too unstable for use in making decisions about individuals. No reliability data are available for the sample of writing ability or the scoring procedure. PATB reliability data are particularly inadequate for females and minorities.

Until adequate evidence exists that a test is sufficiently reliable to be used, it is unwise to use the test as a personnel selection device. To do so violates professional standards for test development and violates EEOC Guidelines because it can lead to false interpretation of test scores.

C. Validity

^{1/}
The evidence for the validity of PATB is not much more encouraging.

^{1/} The validity of a test battery must rest upon a logical relationship between parts of the battery and a comprehensive job analysis or upon a statistical relationship between performance on the tests and performance on the job.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

There is no evidence that comprehensive job analyses have been performed; therefore, no logical basis has been established explicitly for the tests and scales that comprise the PATB. The lack of explicit job analysis data violates standards for selection tests set by the American Psychological Association (APA) and the EEOC.

Of 16 Agency studies relating performance on the tests to performance on the job, only ten provide sufficient information to permit judging the evidence for validity. Even that evidence for validity which is presented in the ten studies is fragmentary, very weak and unconvincing, and does not meet minimum standards set by the APA or the EEOC Guidelines.

Statistical measures of the degree of agreement between test scores and supervisors' ratings of the job performance in particular jobs have been inconsistent and generally low. Measures of the degree of agreement between test scores and similar or identical criteria of job performance, but in different jobs, have also been inconsistent and generally low. Such results are to be expected in the absence of suitably reliable tests.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

✓ There are no validity data of any kind for the writing
✓ sample or for the Strong-Campbell Interest Inventory which
compares an applicant's interest areas with interests of job
incumbents in a wide variety of jobs. Although there have
✓ been some attempts to validate certain items on the Bio-
graphical Information Inventory, evidence for its validity
is lacking.

✓ Four Agency studies have been done on the validity of
PATB test scores for predicting success in foreign language
training. The results of these studies indicate little or no
value of PATB for this purpose.

✓ It is disturbing that the level of confidence with
which results of these studies are reported and used to
predict job performance of applicants greatly exceeds the
level of confidence of the statistical data on PATB validity
and reliability.

D. PATB Narrative Report

Similar reservations apply with respect to the narra-
tive report that is prepared to summarize an applicant's
PATB results. The strong, confident recommendations to hire
or not to hire applicants for specific jobs or in specific

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

components are not supported by the available evidence for validity of the PATB. No written guidelines are available or used by the psychologists in writing the narrative report and the reports tend to vary considerably in unpredictable ways. The variations appear to be due as much or more to the personal idiosyncrasies of the psychologists as to differences in performance on PATB among applicants. For example, the reports of applicants' writing abilities address different criteria and use ambiguous terms to describe those abilities. The scores from the Strong-Campbell Interest Inventory are being consistently misinterpreted by the Psychological Services Staff, Office of Medical Services (PSS/OMS).

Even if reliable and valid tests are available, their potential usefulness is diminished by unreliable or invalid interpretation. Since most of the PATB tests lack adequate reliability and have no demonstrated validity, the narrative reports based on the test scores are misleading and potentially unfair. Recommendations for specific types of employment that are made without adequate validity data promote unfair use of the test results; this practice violates APA standards and EEOC Guidelines.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

E. Relevance for Minorities and Females

There is no evidence that studies of adverse impact as defined by EEOC have been done for PATB or for any other selection procedure used in the Agency. Although no direct evidence of bias or unfairness appears, several factors indicate a serious potential for misuse or unfair use of PATB:

- there is no evidence that minorities were represented in the groups used to norm the tests;
- there is no evidence that minorities have been included in the samples used to determine job-related validity;
- there is evidence that females have not been represented at all in some samples and are underrepresented in other samples;
- the equations being used to make recommendations for hiring employees have been based on inadequate samples of small size and have not been cross-validated. The samples used to generate these equations appear not to have included any minorities and have had very few females;

⁹
~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

-- no reliability data for any of the tests and scales of PATB are available for minorities, and no reliability data for the work attitude scales are available for females.

Only two studies of minority applicants apparently have been done by the Agency. One of these, done in 1974, did not analyze the data correctly and must be disregarded. The other, conducted in 1979, is not conclusive with respect to fairness of the PATB. About the only thing that can be said, based on the data presented, is that PATB does not appear to have any more bias than do the other selection procedures.

Despite the criticism of the PATB, and the inadequacy of validity evidence for it, it is believed that the Agency needs a battery of good selection tests. If testing is eliminated, the only procedures for selecting personnel would be interview, review of past academic records and experience, and personal recommendation. Each of these alternative procedures has major shortcomings, and none can provide the sort of relevant information about an applicant's capabilities that are potentially available from a good selection battery.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

F. IG Assessment

Based on the consultants' analysis of PATB and their evaluation of the Agency's vulnerability to legal challenge, we considered the recommendation that PATB be suspended in its entirety until validity and reliability of its tests have been established. However, sharing the consultants' belief that the Agency needs a battery of good selection tests, and recognizing the possibility that properly-conducted studies may indeed confirm the validity and reliability of some of the PATB tests, we favor a modified course of action.

RECOMMENDATION:

1. The Director of Medical Services discontinue those PATB tests and procedures which the consultants view as indefensible--

- Strong-Campbell Interest Inventory
- Language Aptitude
- Work Attitudes
- Thurstone Temperament Schedule
- Biographical Inventory

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

- Use of prediction equations for recommending job assignments.
- PSS Professional Applicant Testing Report (the narrative report written by OMS/PSS psychologists) in its entirety.

Although the consultants found no suitable evidence for validity of the measures of intellectual ability, and that some of these tests had low reliabilites, we do not recommend that use of these be discontinued at this time. They require modification, and new norms based on a representative sample of current applicants. They also require a logical rationale for their use, based upon a sound job analysis. Pending such developments, they may provide at least some objective measure of intellectual ability. They should be modified, however, with all possible speed, once job analysis data are available to justify their inclusion in the battery.

The consultants identified as the most serious deficiency the absence of job analyses and logical interpretation of the job analysis results upon which to base the selection tests, a violation of APA professional test development standards and EEOC Guidelines. We agree that a logical basis (test rationale) should be derived from

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

a comprehensive job analysis prior to modifying the tests of intellectual ability cited above and prior to adding any measures to the battery to replace the deleted tests. In the interests of expediting this process, initial work in conducting the comprehensive job analysis could be accomplished under contract by a group of cleared test development specialists/consultants, using commercially available instruments. At the same time, the Agency should hire test development specialists for an ongoing capability. We estimate that three to four professional personnel, with supporting staff, would be a workable complement for such a test development unit. Appendix H of this report should be provided to them for initial guidance.

RECOMMENDATIONS:

2. The Deputy Director of Central Intelligence:
 - a. Authorize the Director of Personnel Policy, Planning, and Management to contract with job analysis specialists to assist the the Agency in developing an Agencywide job analysis by a specific date. The Position Analysis Questionnaire, developed at Purdue University, appears to be an appropriate instrument for this purpose.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

- b. Establish a unit of three to four professionals under the Director of Personnel Policy, Planning and Management to develop and try out a new applicant testing program, and establish professionally acceptable reliability and validity data and norms for such tests before they are authorized for administration and use in personnel selection. This unit should also be responsible for assuring the reliability and validity of all other Agency testing for professional and nonprofessional applicants. Job analysts and test development specialists with demonstrated professional training and practical experience in the specialist field of job analysis and test development should be hired for this unit.
- c. Issue a testing policy statement, clarifying which applicants should take selection tests, describing the steps underway to improve the tests, and specifying that component-developed tests are to be validated.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

3. The Director of Medical Services continue to administer the following tests subject to initiating a job analysis and a program to establish reliability data and new norms --

- Vocabulary
- Reading Comprehension
- Figure Matrices
- Arithmetic Reasoning
- Contemporary World Affairs
- Numerical Operations
- Considerations
- Interpretation of Data
- Essay

We agree with the consultants that the results of the eight cognitive tests listed above should routinely be reported in the files of all applicant who are tested, and that names of applicants with special skills and high abilities be retained for computerized recall for an indefinite period. Reports of applicant test results should include the notation that the use of the eight cognitive tests is an interim procedure pending a validated testing program. The essay should be included on an unevaluated basis.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

~~ADMINISTRATIVE-INTERNAL USE ONLY~~

RECOMMENDATIONS:

4. The Director of Medical Services, in consultation with the Director of Personnel Policy, Planning and Management, insure that test results, including the unevaluated essay, are placed in an applicant's file prior to its review by hiring officials.
5. The Director of Personnel Policy, Planning and Management enter into computerized records for indefinite retention the names, special skills and test results of hard-to-get and unusually promising applicants.

~~ADMINISTRATIVE-INTERNAL USE ONLY~~